Chapter 2-Keeping the Brain in Mind

> On its own, a neuron firing has no meaning, no symbolic quality whatsoever.... It is a level shift as drastic as that between molecules and gases that takes place when thought emerges from billions of in-themselves-meaningless neural firings.
>
> —Hofstadter, 1985, p. 649

Virtual reality experiences are making people physically ill. Recently there have been reports of people having flash-back experiences, similar to those associated with LSD. One explanation that has been offered is that when people are harnessed to a virtual reality device, the brain receives visual and auditory signals indicating the body is in motion but it is not receiving the kinesthetic signals that normally go with them. Accordingly, the brain starts adapting to a new environment by establishing new neural pathways, which can then be activated by other signals, thus producing the flash-back phenomena. I have no idea how this conjecture will fare, but the way it is formed is most instructive for thinking about the mind in relation to the brain.

Note that it is the *brain*, not the mind, that ‘expects’ kinesthetic sensations of motion and that starts creating new structures when they do not appear. The *mind*, in one way, is not fooled. We know we are sitting in a room and not behind the wheel of a racing car roaring around a speedway. In another way, the mind is fooled whereas the brain is not. We experience the bodily sensations of movement along with the visual and auditory ones, and so we are unaware of the inconsistencies or error signals that our brains are busy trying to rectify. To describe and make sense of such phenomena, therefore, we need a concept of mind as well as a concept of brain. But the two conceptions ought to be in some accord.

This chapter is about two different models of mind, which have different implications for how the brain relates to mind and knowledge. According to one model, knowledge is encoded in the brain in something like the way that data are encoded in a computer’s memory. This model is fully consistent with folk theory and so it feels comfortable and seems intuitively compelling; but it starts to become less plausible when we begin to trace out its implications at the brain level. According to the other model, the brain does not actually contain knowledge in any sense that we readily conceive of. Thus this model is radically at variance with folk theory and not even comprehensible until we get a handle on how a brain thus constituted could sustain knowledgeable, intelligent behavior. However, it is this second model that I believe we need to develop in order to have a theory of mind that will carry education into the knowledge age.

Computational Models of Mind-as-Container

I wonder whether the pioneers of artificial intelligence ever asked themselves whether they should adopt the mind-as-container metaphor. I doubt it. Simulating human cognition on a computer means writing a program. A program consists of instructions, and these instructions apply to the contents of locations in the computer's memory. Thus, right from the start, we have something that closely resembles the folk conception of mind. There is a container (computer memory) with specified objects in it. These objects are of two kinds: beliefs (data) and rules (instructions). To simulate human cognition, all you have to do is load in data that represent human beliefs, along with instructions that represent the rules human beings follow in operating on those beliefs.

Beliefs can be represented as propositions (e.g., "All birds are bipeds," "Some evangelists are lechers, " "All evangelists are bipeds"). Rules can take the form of logical operators—if-then statements, which have come to be known as "productions." Of course, what must be represented are not ideal beliefs and rules, but rather those that actually appear in human cognition. Thus, to simulate human reasoning, you do not want a flawless logic machine; but you also do not want one that will infer from the preceding propositions that some birds are lechers or that some evangelists are birds. That would be a poor simulation, inasmuch as hardly any human beings would arrive at such conclusions.

Such simulation failures can be avoided in two ways: either by building in more elaborate inferential rules or by building in more extensive knowledge. The first alternative leads to trouble. People do, on occasion, make the kind of logical error represented in "some evangelists are birds," so a valid simulation of human intelligence cannot build in rules that altogether eliminate such errors. The second alternative is more realistic: People avoid error in the present case because they already know that no birds are evangelists. Thus, to simulate human intelligence, it is not enough to build a system that reasons the way we do. The system must also possess what typically turns out to be vast amounts of knowledge that we human beings seem to rely on in dealing with even the most mundane of real world problems.

Work on artificial intelligence has made some notable contributions to our understanding of the mind. If it has not disposed of the mind-body problem, it has at least put its more simplistic versions to rest: Computers, acting as minds, now control all kinds of 'bodies'—from traffic lights to unmanned space vehicles. It has also demonstrated ways to overcome the 'homunculus' problem. In folk theory, the mind is like a homunculus, or little person, who sits in the head, interpreting incoming information and issuing commands to the body. Artificial intelligence programs are homunculi of this sort. The problem arises with the question, "What controls the homunculus?" The homunculus would seem to require a homunculus in its own head, and so on in an infinite regression. A solution that has produced impressive results amounts to endowing the first homunculus with a system of if-then rules that are executed automatically whenever conditions match the 'if' portion of the rule, plus a few simple rules for deciding which of several matching rules gets to go first (Anderson, 1993). This solution may require a proliferation of rules, but it eliminates the need for a higher-level agent to decide which rules to apply when.

From an educational standpoint, however, by far the most important contribution of work on artificial intelligence has been its demonstration of the magnitude of the role of knowledge in cognition (Bereiter & Scardamalia, 1992; Glaser, 1984). It is difficult to convey the revolutionary significance of this result, because it is a matter of degree. Everyone already recognizes that knowledge is essential. Students who have never heard of Abraham Lincoln and the American Civil War cannot be expected to understand the Gettysburg address, no matter how clever and skillful readers they may be. What people do not generally realize is the sheer amount of knowledge involved in mundane intelligent behavior. We notice the special, problematic items of knowledge, but not the ordinary, taken-for-granted ones. We recognize that in order to understand "Four score and seven years ago" you have to know that a score is twenty.

We do not consider that you must also know what a year is and that "ago" indicates years in the past—so that you must also understand time and what it means for something to be in the past; and that 87 years is longer than most people live, so that something that took place 87 years in the past will be outside the experience of the listeners.

In later chapters we will pursue some of the educational implications of this expanded realization of the role of knowledge—not only for ordinary intelligent behavior but more especially for expertise, which, again, depends on previously unrecognized amounts of knowledge, but uncommon rather than common knowledge. For education appropriate to the knowledge age, such realization ought to play a much larger role in educational thought and practice than it does now.

At this point, however, my concern is what this expansive view of the role of knowledge implies for folk theory of mind. Obviously, it implies that the mental container has much more in it—orders of magnitude more—than common sense had suggested. But there are deeper problems. When I said earlier that everyone knows that no birds are evangelists, I cannot have meant that people have that particular fact firmly in mind. If they did, they would also need items specifying every other thing that birds are not, and similarly for everything else of which they have knowledge. That would imply not just a very large number of items in the mental container but a virtually infinite number. So knowledge in the mind cannot be that specific. It must consist of more general propositions or knowledge structures from which, whenever the need arises, we infer things like there being no feathered evangelists.

But do we, in fact, have to infer that no birds are evangelists? Is not this something we simply know, even though we have never thought of it before? Here is where the container metaphor, when carried to the point of an explicit theory or model, becomes implausible. Artificial intelligence models based on the container metaphor offer only two ways for a mind to 'know' something. Either that something is explicitly represented in the mind, in something like a sentence, or else it is logically inferable from things that are explicitly represented. Yet a great deal of what we seem to know does not plausibly belong in either category. We see a china plate hanging over the edge of a table and we push it back. If challenged for a reason, we would respond immediately and confidently that (a) the plate was in danger of falling off and (b) if it fell it would break. The first could be deduced from laws of physics, but that is not something we could do in a flash. The second is not deducible from physical laws, at least not from ones the ordinary person would know. If required to justify the second belief, we might appeal to premises such as 'china plates are fragile,' 'a wood floor is a fairly hard surface,' 'the table is over two feet high,' and 'fragile objects falling from a height of two feet or more on to a hard surface are likely to break.' But would we actually claim to have called up such propositions and performed the necessary logical operations on them? Furthermore, do those propositions represent explicit items of knowledge in our minds or are they also facts that have to be inferred?

Something just isn't right about this way of accounting for human knowledge. To provide a more plausible account, the philosopher Michael Polanyi (1964) invoked the concept of tacit knowledge. The most obvious examples are knowledge embedded in skills—for instance, the knowledge of ballistics implicit in throwing a paper wad into a wastebasket. But, as the preceding example shows, there is also tacit knowledge involved in purely cognitive acts, such as predicting and explaining.

The mere tacitness of knowledge presents no particular problem for folk theory or for theories based on the container metaphor. One simply allows that the mind contains many rules and propositions that the owner is unaware of. No, the problem comes from the intuition that we have knowledge that does not take the form of rules or propositions at all and that this knowledge constitutes a large part—perhaps an overwhelmingly large part—of what we personally know.

That, as I read him, was the intuition Polanyi was trying to win a place for in epistemology. Until about a decade ago, however, that intuition was too insubstantial to make headway against the centuries-long development of formal logic and the dramatic rise of thinking machines, all based on treating knowledge as composed of explicit symbolic representations and thought as the application of definite rules to these representations. As a result, tacit knowledge was regarded as shadowy stuff in the background of the mind, possibly important but not anything you could do much with, compared to the propositions that present themselves in the foreground of the mind, sharp and ready to do work.

All this has begun to change with the rise of connectionism (McClelland, Rumelhart, and the PDP Group, 1986; Rumelhart, McClelland, and the PDP Group, 1986). Connectionism is a different approach to designing artificial intelligence. If the rule-based kind of artificial intelligence I have been discussing up to this point is based on a metaphor of the mind, connectionism may be said to be based on a metaphor of the brain. It is the brain conceived of as a lot of interconnected units, activating or inhibiting each other by energy transmitted over their connections. Connectionism is not a theory. Indeed, from the distant vantage point of education, it serves more as an antitheory. It serves as a source of concrete demonstrations that you can have something like a mind that has something like knowledge, but that does not contain any identifiable rules, propositions, or other symbolic representations of that knowledge.

How Do Brains Know?

When philosophers talk about knowledge, they usually take as examples simple factual propositions (often called "*p*" for short): Kuala Lumpur is the capital of Malaysia. The sun rises in the east. And so on. These are just the kinds of knowledge items that folk theory of mind is best equipped to handle. Imagining that these items are represented as sentences in the mind is not much of a stretch. In fact, it is hard to imagine them not being represented in some sentence-like form. After all, that is often the form in which they were first made known to us. However, just because knowledge of this itemized type seems so unproblemantic, it is a poor source of examples for critical examination of folk theory and alternatives to it.

Let us, therefore, take a different starting point. Consider the following proposition: The world is composed of discrete objects that persist under movement in three-dimensional space. The interesting thing about this item of knowledge is that people show evidence of possessing it in early infancy, and may even be born knowing it (Spelke, 1982). Here there is no question of their having heard it somewhere. If you presented the idea in verbal form to six-year-olds they would have trouble grasping what you were talking about. Yet babies who have not yet uttered their first words will evidence surprise and sometimes alarm if a trickster shows them something that violates the principle: for instance, an object that disappears behind one screen and reappears from behind another, without having been seen to cross the gap between the screens.

An upholder of folk theory could object that what we have here is not a matter of knowledge but a matter of how our perceptual system is built. We are built to see the world in three dimensions and to pick out objects in it and to track those objects as they move about. This tracking of objects presupposes that they persist, but that is not something we know or believe (at least, not as infants); it is, as it were, something our neurons have evolved or been conditioned to expect.

I agree with all of this except for the first statement, that it is not a matter of knowledge. Although it may have been acquired through evolution rather than through learning, it is knowledge very similar to the number sense we discussed earlier or to the acquired knowledge that enables us to find our way about in a familiar environment. To be sure, it is not the same as the knowledge represented by the proposition that the world consists of objects distributed in three-dimensional space. That is a debatable proposition, whereas the knowledge built into our perceptual system is not (although it may profoundly limit our ability to entertain alternatives to the theoretical proposition). But number sense and geographical place knowledge are not debatable either. Propositions based on them may be, but the knowledge guiding our actions is not.

A definitional issue begins to arise here that needs to be cleared up before we can proceed to issues of substance. What counts as knowledge? Gilbert Ryle (1949) is credited with having clarified this question by distinguishing 'knowing-how' from 'knowing-that.' In traditional discussions of epistemology, only the second is counted as knowledge. Knowing-that consists of holding a true or warranted belief. Knowing-how is a matter of skill and is therefore of no interest to epistemology, whose traditional concerns have been with the bases for belief. Cognitive scientists have used the term knowledge much more broadly, however. Knowledge includes false beliefs that the person acts upon as true. But it also includes skills or what has come to be called 'procedural knowledge.' One reason for lumping all these together as knowledge is that, on close inspection of cognitive behavior, the separation between knowing-how and knowing-that is not very firm. Today you may recall and therefore be said to 'know that' the square of 16 is 256. Another day you may not recall it, or may recall it but not be sure it is right, and so will rely on your ‘knowing how’ to multiply and thereby arrive at an answer. When we turn to more complex sorts of competence, the distinction becomes even less tenable. Few people know explicit rules for transforming English declarative sentences into questions and even linguists find some contingencies hard to account for. But native speakers can all make such transformations. Thus they could be credited with knowing-how but not knowing-that. Native speakers can do more than this, however. They can also immediately detect nonstandard transformations of the kinds foreigners may make, such as “What you are doing?” or “How much it does cost?” In other words, they ‘know that’ such sentences are not standard English but they do not ‘know how’ to explain what is wrong with them. Thus knowing-how and knowing-that seem to be aspects of the same knowledge, where knowledge of language is concerned.

Once we allow that knowledge consists of more than statable beliefs, it is not clear where to draw the line after that. This uncertainty of delineation is evident in ordinary speech. People talk about their body’s knowing when it needs a rest or about what their muscles know. It is never clear where literal meaning stops and figurative begins.

I propose that this is not merely a definitional matter but is a matter of what knowledge really amounts to at the level of the brain. ‘Really,’ I suppose, should be in scare quotes, because I do not want to have to defend use of the term. I merely use it as a reminder that, although we may be free to define knowledge and the mind any way we like, there is also a brain involved, and we do not enjoy nearly so much freedom in what we may attribute to it. You may or may not choose to attribute knowledge to the brain, but something or other characterizes the brain and is the counterpart of our ‘knowing that *p*’ or knowing Spanish or whatever. To avoid prejudging issues, let us say that *the brain supports knowledgeable action*. How it does this will be our concern in this chapter.

There are two major contenders as answers to how the brain supports knowledgeable action. This is a big step up from the state of things before 1986, when there was only one major contender—as Jerry Fodor put it, “the only game in town” (Fodor, 1985, p. 90). The older of the contending answers comes from folk theory, as systematized by cognitive scientists. It is that the brain contains encoded versions of propositions and rules. How the brain does this is a problem for a different science. Suffice it to say that with billions of neurons, countless synapses connecting them, and new synapses being created all the time, there are plenty of degrees of freedom available for encoding all the knowledge that could accumulate in the course of an individual lifetime.

The nature of this code is a problem for cognitive science, however, and ultimately for education and other applied fields that must deal with mental states and personal knowledge. The brain's code cannot be like the Morse code and other such codes or like a scrambling device. These merely encode the sounds or spellings of linguistic utterances. That would leave the prelinguistic infant without a mind and would mean that "some dogs bark" and "there exist dogs that bark" are separate items in the mental storehouse. No, for this theory to work, the brain must encode concepts, meanings. Thus, it must have a grammar and a vocabulary. In short, the brain must have a language (Fodor, 1975).

This is just what philosophers and others who have thought about the matter under the influence of the container metaphor have assumed all along. Fodor's distinctive contribution has been to pursue this assumption far enough to reveal its more profound implications for a theory of mind—or, some would say, far enough to reveal its utter implausibility. This 'language of thought' as Fodor called it in his 1975 book, has to be innate. Otherwise the infant would have nothing to get started with in learning. Thus the brain has to come already furnished with a grammar—as Chomsky (1975), on other grounds, had been arguing that it must—and with at least some initial vocabulary. Fodor went on to argue, however, that this initial endowment of language cannot be merely some neural version of baby talk. It has to embody the highest-level concepts and structures that the mature mind will exhibit. In effect, the learning of more complex cognitive structures is impossible, because it would involve using the existing language of thought to create a language capable of representing knowledge beyond what can be conceived of in terms of the existing language:

> There literally isn't such a thing as the notion of learning a conceptual system richer than the one that one already has; we simply have no idea of what it would be like to get from a conceptually impoverished to a conceptually richer system by anything like a process of learning (Fodor, 1980, p. 149).

Such a paradoxical view of learning could not go unchallenged, and there have been challenges ranging from metatheoretical appeals to self-organizing systems (Molenaar, 1986) to arguments for developmental processes distinct from learning (Boom, 1991). But all of these by-pass the problem of showing how a rule-based system can generate rules of a higher order than those that do the generating (Bereiter, 1985, 1991a). This is the challenge taken up by the second contender, connectionism.

Connectionism: Mind Without Mental Content

The question, as I put it earlier, is how does the brain support knowledgeable action? The reason for phrasing the question in this quaint way is in order to leave open the possibility that the brain can support knowledgeable action without actually containing the rules, propositions, images, recorded events, and so on, that are conventionally reckoned to constitute knowledge. That is the kind of situation that connectionist and neural net models simulate.[1]

Before proceeding, I need to try to clarify this notion of containing or not containing rules. Anything that works can be said to embody rules. A mechanical clock embodies a rule, enforced by the pendulum or balance wheel, that the minute hand shall go around once and only once per hour. A doorbell circuit embodies the rule that the doorbell sounds only when the button is pushed. As simple a device as a pair of pliers embodies the rather formidable rule that the force exerted by the jaws is equal to the force of the grip, multiplied by the ratio of two relevant distances from the pivot. Naturally the brain embodies rules of this sort, but they are electrochemical rules and the like. Whether, in addition to that and in quite a different sense, the brain contains rules for making butter tarts and for pleading cases in court is the question before us and one not to be confused by characteristics that the brain shares with every other mechanism under the sun.[2]

The computer programs that implement connectionist models, being things that work, naturally also embody rules. These, however, are mathematical rules governing how the program runs, and they bear no relation to rules of the domain that is being modeled. A connectionist network consists of a (sometimes large) number of units with connections between them. Some are input units, which receive quantitative values from outside. Some are output units, which output quantitative values that are interpreted and sometimes fed back to the input units. The rest are ‘hidden’ units whose inputs come from other units in the network and whose outputs similarly go only to other units in the network. Although in principle the whole system is active at once, connectionist programs that run on conventional computers simulate parallel processing by going through cycles. In each cycle, quantitative outputs pass from unit to unit, via the connections, changing the states of units that receive them so that they may output different values on the next cycle. The rules, such as they are, determine these quantitative values and the effects they will have. The rules include parameters that the operator can adjust to 'tune' the network. When properly tuned, the network will, after going through enough cycles, 'settle' into a state such that the inputs that units receive on each cycle leave them in approximately the same state they were in before, so that the outputs no longer change.

As described, this is as mindless a system as we could ask for. Thus it models, at least abstractly, the paradoxical condition Hofstadter referred to in saying, "On its own, a neuron firing has no meaning, no symbolic quality whatsoever." If a system composed of such meaningless elements and processes can begin to do meaningful, knowledgeable things, that is at least interesting and possibly illuminating as to our question, "How does the brain support knowledgeable action?"

Connectionist networks have learned to do meaningful, knowledgeable things. Many of them are not very impressive by human standards, because they are things people do easily: recognizing letters of the alphabet, guiding a car on a real road, learning grammatical conventions, recognizing faces. These are impressive by artificial intelligence standards, however, because they are accomplishments that have proved inordinately difficult for rule-based systems. Others, such as learning the evasive maneuvers of a fighter pilot and making insightful investment decisions, are impressive even by human standards. In all cases, one would look in vain in the program code or in its output for the rules that the network is following in accomplishing these things. That is because there are no rules in that sense. Connectionist networks can act knowledgeably without containing knowledge; they can behave lawfully without containing rules.

Educators who have gathered this much about connectionist models recognize, quite properly, that they do not have much of practical value to contribute to pedagogy. (I ignore here the value connectionist technology may have in educational software applications [cf. Schank, Ranney, Hoadley, Diehl, & Neff, 1994] or in understanding neurological factors in learning difficulties [Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996]. These are potentially important, but not germane to present issues.) That is, there is not much to be learned from a system that can do something intelligent but cannot tell you how it did it. Rule-based systems win easily on this account. A rule-based system that can read and solve mathematical word problems, for instance, contains a set of rules that are potentially teachable to students. This kind of potential was in fact exploited by Clancey (1987), who took an expert system for medical diagnosis and used it as the basis for a tutoring system that would teach medical students the propositional knowledge and the reasoning procedures the system used in making its diagnoses. Nothing like that can be expected from connectionist research. Instead, the whole reason I have been trying to disseminate connectionist ideas among educators has been to provide a handle on a new way of thinking about knowledge and the mind.

To gain such a handle, one needs to acquire at least some experiential feel for how connectionist networks work. By far the best way is to get hold of some simple programs and fool around with them on a computer. Some of the pioneers of connectionism have made this easy by producing a handbook that comes with software in DOS or Macintosh versions (McClelland & Rumelhart, 1988). Most writings on the subject are at too abstract a level to give much feel. I have tried to provide something more down-to-earth using a network of frisbees as an analogy (Bereiter, 1991b). Here I will take a different down-to-earth approach, building on an example that will be familiar to people who use word processors.

The Mind of the Spelling Checker

The kinds of spelling checkers now found in most word processors are just barely complex enough to serve as examples for comparing the two models of mind we have been discussing. As far as I know, the spelling checkers in actual use are all built according to the rule-based model. That is, their source code contains explicit rules for checking spelling. Let us pretend, however, that we do not know this and that we are free to speculate about the mind of the spelling checker, constrained only by what we are able to observe of its behavior.

Starting at the behavioral level, this is what we observe: The spelling checker proceeds through a text sequentially. From time to time it highlights a word and presents a box giving us the option of changing the word or going on, with or without adding the highlighted word to the dictionary. It may also present a list of alternatives, words similar in spelling to the one highlighted, among which we may find correctly spelled the word we intended. In that case we may select the alternative and it will replace the highlighted word.

A mentalistic description captures more of the point of what goes on. The spelling checker *searches* through the text. When it *finds* a word it does not *recognize*, it opens a *dialog* with us. It *suggests* alternative spellings--we might even say it *tries to guess* what word we had in mind. And so on. The description raises further questions of a psychological sort. An interesting one is how the spelling checker *decides* which alternatives to suggest. We could begin to investigate this question by experimentation. The spelling checker I am using does puzzling things with variations on "psychology." Here are examples:

Word in Text	Spelling Checker's Suggestions
pyschology	psychology, psychologies, psychologic, psychological
pyschilogy, pyschalogy	psychology, psychologic
pyschiology	no suggestion
dsychology	no suggestion
sychology	psychology, psychologies, psychologic, psychological
cychology	cytology

It seems that first letters carry quite a bit of weight—yet the spelling checker is not thrown off by "sychology." Perhaps the number of letters in common determines a suggestion—but then why does it give up on "pyschiology"? I expected it might bring up both "psychology" and "physiology" as suggestions[3]....

What am I doing in these speculations? I am doing folk psychology, applied to an inanimate piece of software. I am trying to figure out what is in the spelling checker's mind, what rules it follows in arriving at judgments and decisions. Besides noticing what suggestions it gives in response to various misspellings, I also noticed that it seems to ‘think’ longer about some misspellings than others. For instance, it takes a longer time to respond to "sychology" than to "cychology." That, too, is information I might be able to use in forming a theory of its behavior. All this is not much different from what a teacher might do in trying to make sense of a child's idiosyncratic spelling habits.

The spelling checker can also be described at lower levels, which correspond to the neurological level in humans. It can be described at the level of subroutines, of source code, and of binary code. But even the programmers who understand the spelling checker at that level will still have reason to think about it at a mentalistic level. If they want to improve the spelling checker so that it does not give up on so many words and so that it recognizes "cychology" as a plausible misspelling of "psychology," they will not start by messing around with the code. They will have to start by, in effect, thinking like a teacher: What could I teach this spelling checker so that it would do a better job? Only after they have worked matters out at this level can they do what is denied to the teacher of living organisms: get inside and directly alter things in the mind.

The mental level of description does not, however, require a commitment to folk theory of mind. It does not, in other words, require us to believe that those rules we have been inferring are actually inside there somehow and determining what the spelling checker does. It could be that the spelling checker merely behaves in ways that correspond approximately to the rules we have thought of. The fact that our made-up rules do not fully account for the spelling checker’s behavior may not be a fault of those rules. It may be that *no* rules could fully account for its behavior. In the case of a normal spelling checker, we know, of course, that there is a right set of rules; these are the rules that the software developer built into the program and, barring bugs, they will fully account for its behavior. But if we did not have this outside information, we could not be sure. If all we had was observed behavior, we could not rule out the possibility that we were dealing with a connectionist system that contained no explicit rules.

A connectionist version of a spelling checker might be built as follows: Initially, some simple method is used to compute similarities between words in the dictionary. These are stored as connection weights between units that represent dictionary words.[4] There are also a number of other units that have connection weights to each other and to word units. Some of these are feature units. One feature might be "starts with the letter 'p.'" This unit would have positive connection weights to every word that starts with "p" and negative weights to every other word. A number of other units would have no assigned identify at all, although there would be arbitrary connection weights to the other units. When the spelling checker is asked to suggest alternatives to a text word, feature units corresponding to the text word would be activated. These would send activation to other units, according to the connection weights, these units would send activation to others, and so on. Activation, positive or negative, would be sent back to the feature units, thus leading to another cycle of activation, and so on. The network of connections is designed so that with repeated cycles the activation levels tend to stabilize. When a point of stability is reached, the word units having the highest activation levels are chosen as suggested alternatives.

At first, such a system would probably not perform as well as the rule-based systems now in use. But it can be designed so as to learn. Every time a suggested word is selected as an alternative, the connection weights leading to it are increased a little, as are the weights leading to the units that sent positive activation to the units that activated the word. Although it might seem that this would produce nothing more than a tendency to repeat past successes, experience with learning systems of this kind indicates that the learning can be much richer than that. A well-designed and well-trained system might be expected to match up words on the basis of pronounciation as well as matching letter sequences and, indeed, to work out useful correspondences that the designers had never thought of. The spelling checker would seem to be working according to rules--sometimes obvious ones, other times rules too subtle for an observer to formulate. But this would be an illusion.

The same possibilities apply to the human mind. When we explain behavior using the conceptual framework of folk theory of mind—that is, when we explain behavior in terms of beliefs, plans, desires, and the like—we speak as if these things were actually in the mind of the behaver. We assume they are encoded in the brain somehow, and that if we had access to the code and could understand it, could interpret the molecular structures or the patterns of synapses, we could know what a person's real beliefs and plans are. But the alternative is that the mind does not actually contain the entities folk theory of mind posits, that the brain is like a connectionist machine, producing rule-like behavior without actually containing such rules in any representation whatever. The fact that the theories we construct on the basis of folk theory of mind work in predicting behavior is no proof of the first alternative. But the fact that the brain consists of billions of neurons with billions of interconnections is no proof of the second alternative, either. Those billions of synaptic connections might or might not encode what folk theory of mind assumes they do, and it is not obvious how studying the brain itself could tell us which is true.

"This whole argument ignores consciousness," you might object. "We can't be sure what goes on in the so-called mind of the spelling checker, because the spelling checker has no way of telling us. But I am conscious of at least some of what goes on in my mind, and I know that my mind contains beliefs and rules and so on, that it is not just a lot of meaningless connections." But let's test this claim out. What does your consciousness tell you about rules versus connection weights when it comes to checking visually for spelling errors? Ignore "*i* before *e* except after *c*" and a few other explicit rules that you may know, for they obviously account for only a tiny part of what goes on when you are checking for spelling errors. What I am mainly conscious of in proofreading a manuscript is running my eyes over the words and having nothing happen. I am certainly not conscious of applying rules to the words or of comparing each word to models stored in my mind. From time to time I spot an error and immediately correct it. For instance, I notice and delete the superfluous 't' in 'intutition.' Here, again, I am not aware of any rule application or indeed of any thinking at all. I simply notice and act. But then there are times when I look at a word and think, "This just doesn't look right." If I do not have an authoritative source at hand, I may try several different spellings of the word and hope that one of them clicks as looking right. All of this sounds to me much more like a connectionist system than a rule-based one. I could be wrong. Maybe my brain contains 200 or more spelling rules[5] and runs through them at lightning speed; but my conscious awareness is no help in deciding if this is the case.

The Mind and the World

The connectionist mind is much more directly in contact with the world than is the symbol-processing mind conferred upon us by folk psychology. Being ‘more directly in contact with the world’ does not, of course, mean bypassing sense organs and neural pathways. Remember, it is the mind, not the brain we are talking about, and so the directness of contact is psychological. I like to think of psychological directness along the lines that Joseph Church (1961), a phenomenologist, talked about the experience of writing with a pencil on rough paper. The sensation comes to us through vibrations of the pencil moving over the paper, but we do not experience the vibrations of the pencil. Instead, we feel the texture of the paper directly. Similarly, according to the connectionist view, we do not receive a pattern of visual stimuli which we interpret or recognize as a bear. Our visual system sees things, because that is what it is built to do. Given certain inputs, we will see a bear. There is similar directness with respect to action. In the symbol-processing mind, the belief that a bear is approaching activates fear and a variety of escape schemata, while the connectionist mind has us clambering up the nearest tree.

Putting it more generally, the connectionist view of mind is at its best in making sense of people interacting in real time with things in the external world. The symbol-processing view interposes too many unlikely mental events. The symbol-processing view is best suited to rationalizing such behavior after the fact—to explaining, for instance, how one might reason one’s way into climbing a tree to escape from a bear. In this case, the rationalized version probably has little resemblance to the actual mental events, which are likely to have been a jumble of thoughts and impressions accompanying rather than causing the physical actions. How do we explain someone’s climbing a tree to escape a bear when they know that bears can climb trees? For those who adopt the symbol-processing view, such behavior raises questions about human rationality. But from a connectionist standpoint, there is nothing paradoxical about it. Given the real-world situation, there are only a few alternatives on which the cognitive system would have much likelihood of settling: confront the bear; cower; turn tail and run (with the prospect of being seized from behind); step aside in hopes the bear will go past; and climb the tree. Assuming there was no strong bias from prior experience, which option prevailed would likely depend on particularities of the situation. The sight of a reachable branch could well tip the balance in favor of tree climbing, whereas a slightly higher branch or a slightly less menacing bear might tip it in the direction of one of the other possibilities. Similarly, having previously seen a bear climb a tree (as opposed to having only heard that bears can do so) would no doubt greatly lower the attractiveness of the tree-climbing alternative.

There are cases, however, in which the rationalized version corresponds more closely to the experienced stream of thought. These are cases in which we are trying to work our way deliberately, step-by-step through a problem or a procedure. On such occasions we do consciously invoke rules, recall facts, draw conclusions, question a belief. Indeed, when we are conscious of thinking, those are the kinds of things we are conscious of. No one is questioning that we have such abilities or that the experiences are genuine. There are times when we do act like logic machines, applying mental operations to symbolic representations.

Some cognitive scientists opt for a division of labor, awarding to a connectionist part of the mind the job of fleeing from bears and to a symbol-processing part the job of figuring out why going up a tree is not such a good idea. Keil and Silberstein (1996) argue very persuasively for what amount to two worlds inside the head, one part being occupied with theorizing and reasoning on the basis of causal relationships, the other part (the connectionist part) building up competence based on covariances (events and conditions that statistically go together in the environment, whether causally related or not). Others, such as John Anderson (1993) and Walter Kintsch (1998), treat it as a matter of levels. There is a cognitive level that works with rules or propositional representations and a subcognitive or subsymbolic level that functions in a connectionist manner. But a number of the world's cleverest people are busy trying to build connectionist systems that do the logical things that skeptics maintain they cannot do. To the extent that they succeed, we shall have support for a conception of the mind that is connectionist all the way to the top. According to this view, there are no statements or rules for manipulating them in the head, it is just that, under favorable conditions, human beings are capable of doing a pretty good job of simulating the kind of logic machine that does manipulate stored symbols according to rules.

No matter how the theoretical differences are eventually worked out, there is one important departure from folk psychology that seems already too well established to fail any time soon. Because we are most conscious of mental activity at those times when we are deliberately applying rules to propositions, folk psychology assumes that the same kind of activity is going on all the rest of the time, even though we are not aware of it or are not paying attention. The alternative view, urged on us by both connectionists and situativity theorists (e.g., Suchman, 1987), is that conscious thought is not simply being conscious of something that otherwise goes on unconsciously, it is a special kind of activity. We are not bystanders observing the flow of thought. Paying attention is a crucial part of thought itself. It produces a kind of thinking that does not occur otherwise, thinking that is systematic and deductive. This kind of thinking does not come easily to us. It is severely limited by the number of things we can pay attention to at the same time (Case, 1985; Miller, 1956), and it is difficult to sustain over a long series of steps unless we have external aids. It has obvious strengths, as perhaps best exemplified by applied mathematics. Faced with a real-world problem, we can sometimes create a mathematical representation of it and then, by puzzling our way through the mathematical problem, arrive at a solution which, lo and behold! translates into a solution to the real-world problem. But most people go through 10 years or more of school mathematics without ever acquiring much facility in performing such a feat. Their normal ways of dealing with quantitative problems rely much more on affordances of the situation and on number sense (Lave, 1988), exhibiting behavior that is better modeled by connectionist than by rule-based systems.

Mind and Emotion

When asked about the role of emotions in cognition, an eminent cognitive scientist dismissed the question, saying, “That’s just a hardware problem.” In effect, he was saying that emotions may affect how well the cognitive computer runs its programs, but they have nothing to do with the programs themselves. Of course, there is some truth to that. Emotions can make the mind go haywire. But both common sense and decades of research on emotions and motivation argue for emotions having a much more extensive and complex role in our mental lives than that. These sources of wisdom argue, furthermore, that emotions can be beneficial as well as detrimental. From an evolutionary standpoint it can be argued that emotions must be good for something, otherwise we wouldn’t have them. The unemotional would long ago have inherited the earth. However, it can also be argued that emotional volatility is a price we pay for having such a powerful and finely tuned cognitive engine.[6]

The problem for folk psychology and its symbol-processing derivatives is how to give emotions a role *in* cognition as contrasted with a relationship in which emotions and thoughts have effects *on* each another but are otherwise as distinct as the brain and the stomach—which also influence each other in sometimes quite pronounced ways. If the mind consists of mental content and rules for operating on it, emotions have no obvious place. They are not rules and they are not content in the usual sense. Emotions are states. That is how common sense puts it, in saying that an emotionally upset person is "in a state," and neuroscience offers us no reason to characterize emotions differently. How, then, can these states—which are physical as well as mental—affect cognition? They cannot function as rules and they function as content only to the limited extent that we can have knowledge of our emotional states and use that knowledge in reaching decisions ("I'm too angry to talk calmly with him now; let's put off the meeting till tomorrow").

The prevailing view among cognitive psychologists who study emotion seems to be that emotions influence cognition by affecting attention and priorities (Oatley, 1992). Fear is not only a response to danger, it attunes us to look for danger. Anger sensitizes us to affronts. When we feel depressed we notice all the things that are wrong with ourselves and our lives. The result can be harmful build-ups of emotion—acute or chronic anxiety, rage, or depression—but it can also lead to appropriate mobilization of our resources to meet problems and opportunities. Some clinical psychologists have proposed that depression, when it is brought on by things going badly, has adaptive value because it causes us to suspend action and think about what is wrong. Similarly, positive emotions of affection and happiness direct our attention to positive things that offer prospect of enhancing our well-being (and, in turn, that of the people we feel positively toward).

One of the problems for rule-based models of cognition is that at any given time there are likely to be a number of rules ready to be activated. (Rules have the form of if-then or condition-action pairings, and so whenever the 'if' conditions are met, the rule is ready to fire.) Thus there has to be an auxilliary process or set of rules that resolve issues of precedence (Anderson, 1983). Emotions could play this role, boosting the priority of certain rules and reducing that of others. They would not likely impinge on individual rules but could affect the priority of combinations of rules, such as plans. According to Oatley and Johnson-Laird (1987), emotions play a monitoring function with respect to people's goals. When something happens to alter the probability of achieving some goal, an emotional response

> draws the attention of the rest of the cognitive system, and sometimes of other individuals, to the kind of goal-relevant event that has occurred, a subgoal achievement, a goal loss, a goal conflict, or the like. The control signal tends to set the cognitive system into a distinctive mode appropriate to this kind of event. It makes ready a certain reportoire of actions, and it creates a preoccupation with what has occurred that assists further planning. (Oatley, 1992, p. 174).

These are important functions of emotions, but they still leave emotions on the outside of cognition, acting as a sort of Greek chorus, calling attention to discrepancies, cheering and encouraging one kind of action, discouraging another, but not contributing substance of their own to cognitive processes. In connectionist models, however, emotions enjoy a much more natural fit. Emotions are mental states, but so are thoughts and knowledge. In principle they need not be separated at all. Instead of thoughts arousing emotions and emotions triggering thoughts, we may simply think of feelings and cognitive content as attributes of any state of knowing. Often the content is vivid but the feelings are negligible, as when we learn that the square of 15 is 225. Sometimes both the content and the feelings are prominent, as when we receive alarming or cheering news. But of special interest are feelings about objects of knowledge accompanied by little or no content. These are instances when we say about an idea or a situation that it "feels good" or "doesn't feel right" but can say nothing much beyond that. Yet such feelings function as knowledge. They serve as guides to action and they play a central role in creative thinking (Bereiter & Scardamalia, 1993, Ch. 5). In ordinary language they are referred to as intuitions and are mysterious. For symbol-processing conceptions of mind they are bound to remain mysterious, because they cannot be explained in terms of mental operations performed on mental objects. But for connectionist conceptions of mind there is nothing mysterious about them at all. Feeling good about an idea is a knowledge state as much as any other. It is what I will discuss in Chapter 5 as 'impressionistic knowledge,' but it is not categorically different from other knowledge. It lies at one end of a continuum, with things like knowing the square of 15 at the other end, and infinite gradations in between.

The Rediscovery of Human Nature

Let us return to the matter of babies apparently being born knowing that the world is composed of objects distributed in three-dimensional space. Research on young children has given reason to believe that babies are born either already endowed with a considerable amount of knowledge or at least biologically prepared so that they will readily acquire knowledge that takes certain predetermined forms (Hirschfeld & Gelman, 1994). Instead of the "booming, buzzing confusion" that William James said babies are born into, it seems that they are born into a world that their brains have already organized for them. Besides the general spatial organization just referred to, which enables them quickly to pick out and learn to identify objects, they are especially attuned to pick out one particular kind of object: the human face. Besides being disposed to associate sounds with their sources, they are especially attentive to linguistic sounds. The speed with which they master the awesome complexity of language suggests that their brains automatically organize speech input in ways that conform to linguistic structures. Children also appear to grasp number quickly and in ways that go beyond what they could have gained from direct experience or instruction. For instance, at a time when they can only count up to ten, they seem already to have grasped the idea of infinity—that numbers go on increasing indefinitely, that there is always a next number. The infant's mental world seems to be organized causally: Young children spontaneously seek out causes for events and predict and watch for consequences. Along with this causal orientation comes a distinction between animate and inanimate things and a disposition to assign 'essences' to things—inner characteristics that remain constant despite outward changes and that determine how things will behave.[7]

From studies of identical twins come indications that the innate mental endowment may include many more specific dispositions as well. By now most people will have heard the amazing tales of twins separated as infants who are brought together in middle age and find that they have the same breed of dog as a pet and have given it the same name, that they have the same preferences in food, music, politics, sexual partners, and practically everything else. One need take only a small part of this seriously in order to appreciate that theories of mind face a significant challenge in accounting for what our brains already have in them at birth.

Ideas that link genetics and psychology inevitably encounter heated opposition. Ever since the Hitler era there has been the fear that accepting such links will open the door to eugenics and racial cleansing. But there is an additional obstacle to accepting the ideas summarized here, and that is that folk theory of mind makes them seem utterly mysterious. If the genes contain instructions that determine our fingerprints, such that identical twins have identical fingerprints, they could certainly contain an instruction to buy an Oldsmobile or a set of rules for parsing sentences. That is not the problem. The problem is how knowledge, attitudes, and the like that people acquire through experience could get translated into genetic material that gets passed on to their offspring. That sounds like inheritance of acquired characteristics, which we have all been taught is wrong. Invoking natural selection does not solve the problem, however. Being able to recognize human faces no doubt has survival value for the infant, but how does the *knowledge* that distinguishes a human face from the same set of features arranged differently get from the mind into the genes? And how does it then get from the genes into the inheritor’s mind? As far as I can see, these are unanswerable questions as long as we hold to a theory that posits the brain as one thing and mental content as another.

From a connectionist standpoint, there is nothing much more puzzling about the inheritance of knowledge than about the inheritance of any other complex characteristic. We are born with a connectionist network of sorts in our heads that is already running. As we learn, the network undergoes modifications. But what we learn and how we learn are heavily influenced by the initial state and structure of the network, which in turn are largely determined by the genes. At different points in its history, a connectionist network may be settled in its behavior, responding to inputs in stable and highly differentiated ways, or it may be quite unsettled, responding to inputs in variable and and only grossly differentiated ways. It is still the same network. It is not a matter of there being different content or different amounts of content at different times.

Throughout most of its brief history, educational psychology has been absorbed with individual differences, and nature-nurture issues have been considered only from that standpoint. Attention to the genetic side yielded little of pedagogical interest, being an aspect of individual differences that education could not be expected to do anything about. As Benjamin Bloom (1969, p. 419) put it, “The educator must be an environmentalist.... If heredity imposes a limit—so be it. The educator must work with what is left, whether it be 20 percent of the variance or 50 percent.” At the same time, however, pedagogical wisdom put great stress on ‘starting where the child is at’—that is, building on students’ existing knowledge and skills instead of starting from zero or some arbitrary point, as prescribed curricula are wont to do. It is now becoming evident that applying this wisdom is a much more complex business than our default environmentalism had led us to imagine.

‘Where the child is at’ and what the child is best disposed to learn are turning out to have a great deal to do with ‘nature.’ But it is not nature conceived of as inherited individual differences. It is nature conceived of as the whole complex of cognitive attunements, dispositions, and prestructurings of knowledge that characterize our species. For instance, there is some reason to believe that we are born with a network well attuned to additive numerical structures, such as the structure represented by the whole number line, but that we have no prior attunement to multiplicative structures, such as we encounter with ratios and proportions, and that this goes some way toward explaining why ratios and proportions are so hard to learn (Resnick, 1987).

Connectionism is sometimes referred to as ‘the new connectionism’ to distinguish it from an obsolete theory of the same name advanced by E. L. Thorndike (1949) and traceable back to John Locke and British ‘associationism.’[8] Now we have a ‘new innatism,’ different in focus and intent from the older kind. In combination, these provide a novel perspective on what a new theory of mind for the knowledge age should be like. Rather than its being some kind of high-tech theory that conceives of the mind at a very abstract level of logic or knowledge processing, it must be a theory rooted in the fact that we are biological organisms—in short, animals. Like all chordates, we come into the world with nervous systems designed to perceive the world, relate to it, and act toward it in certain ways. We are animals with an amazing range of capacities, which include a capacity to create and share bodies of abstract knowledge. But we do this using a brain whose design evolved under conditions radically different from those of a 21st-century knowledge worker.

> The key to understanding how the modern mind works is to realize that its circuits were not designed to solve the day-to-day problems of a modern American—they were designed to solve the day-to-day problems of our hunter-gatherer ancestors. These stone age priorities produced a brain far better at solving some problems than others. For example, it is easier for us to deal with small, hunter-gatherer-band sized groups of people than with crowds of thousands; it is easier for us to learn to fear snakes than electric sockets, even though electric sockets pose a larger threat than snakes do in most American communities. In many cases, our brains are better at solving the kinds of problems our ancestors faced on the African savannahs than they are at solving the more familiar tasks we face in a college classroom or a modern city. In saying that our modern skulls house a stone age mind, we do not mean to imply that our minds are unsophisticated. Quite the contrary: they are very sophisticated computers, whose circuits are elegantly designed to solve the kinds of problems our ancestors routinely faced. (Cosmides & Tooby, 1999)

How that stone age brain biases and constrains the knowledge we create—including the knowledge we create about ourselves—is something that developmental psychologists are only beginning to reckon with. At this time it is impossible to be definite about what the educational implications of rediscovering human nature will be. Some general expectations are these:

• Curricula may be strengthened by building on innate knowledge or knowledge-constructing tendencies. This is already happening in elementary mathematics instruction.

• As knowledge advances in all the disciplines, it seems likely that more and more of what is to be learned will be counterintuitive in the way that theories in physics tend to be. Accordingly, it will become increasingly important to recognize obstacles imposed by built-in ways of thinking and conceptualizing.

• The kinds of behavior and dispositions required in educational institutions may run contrary to innate dispositions. For instance, distractability—now more ominously labeled 'attentional deficit'—has come to be recognized as a trait, found in a substantial minority of children, which interferes with school learning. Yet if we consider a group of our ancestors engaged in hunting or gathering in a dangerous environment, a good arrangement would be to have most individuals in the group steadily focused on their tasks, while a significant minority are highly sensitive to extraneous stimuli that may signal danger or an unanticipated opportunity. Thus, what we see in classrooms may represent a generally beneficial species characteristic that happens to be ill-adapted to a particular recently-created environment. This is only a speculation, but it illustrates the relevance of an evolutionary perspective on contemporary behavioral problems.

These are all in the nature of promissory notes. We do not know that recognizing built-in resistances to advanced learning will lead to our being able to do anything about those resistances. An evolutionary perspective on behavioral and attitudinal problems may have explanatory value yet not turn out to have practical value. But that is always how it is with advances in theoretical sciences. Practical people have to bet on futures. For those who sense that the burgeoning research on human nature will have future payoffs, there is good reason to start looking for a theory of mind that can make sense of this research.

A Commonsense Way of Going Beyond Common Sense

What I have been calling the 'connectionist' mind could as well have been called the 'self-organizing' mind. Connectionism is part of a much larger movement, affecting all the sciences, toward trying to provide rigorous explanations of emergent phenomena. Self-organization is one of the key ideas. Considering the rate at which the field is developing, connectionist models as they are presently known are sure to become old hat very soon, and the name itself is likely to be abandoned when its associations become bothersome. I am going to keep using the terms 'connectionist' and 'connectionism,' because of their currency[9]; but what really matters is not the particular scientific models that go by that name, it is the way of thinking that they represent.

Wherever biological or behavioral scientists look these days, they see marvelously complex systems that represent an assembly of much simpler components that have somehow got themselves organized without benefit of a higher, guiding intelligence. This is the case with evolution (Dennett, 1995; Kauffman, 1993), termite colonies (Wilson, 1996), and child development (Thelen & Smith, 1994). Human learning itself must have this character; otherwise it is impossible to explain how we can ever learn anything more complex than what we already know (Quartz, 1993). In the popular and educational media there is overheated talk about fractals, chaos, and how the fluttering of a butterfly's wing can affect the climate on the other side of the world; but these effusions should not obscure the fact that a lot of serious work is underway.

A new level of explanation seems to be emerging. We all know the shortcomings of explanations that take the form A causes B, B causes C, and so on in a chain. A second level of explanation takes the form A, B, C, etc., *in combination* produce X. You can find everything from creative genius to heart disease explained in this way. The weakness of this kind of explanation is that it offers no explanation of how the various factors act to produce a result. It is just empirical. The new level of explanation depends heavily on computer simulation (which is partly why it is new). Simulate A, B, and C by creating computational objects that behave like A, B, and C. Also simulate the environment in which they act. Then set A, B, and C and their confederates in motion within the virtual environment and demonstrate that X emerges. With success, you have not only improved the prediction of X but have demystified it. You have shown how it comes about through the interaction of processes (the behavior of A, B, and C) that were already understood. Connectionist simulations have this character. Feed in data about the frequency with which various objects are found in the same room and you witness the emergence of concepts like *kitchen* and *bedroom*, such that the system can now make reasonable inferences about what room it is, given information about objects in it, or about what a room contains, given its type (Rumelhart, Smolensky, McClelland, & Hinton, 1986). Yet no definitions or rules have been introduced and none are to be found in the system. The concepts of *kitchen* and so on are the result of self-organization among the simpler elements.

Common sense is not well equipped in general to think along the lines of self-organizing systems, but it is especially handicapped in thinking this way about learning. Some version of A causes B or A, in conjunction with B, causes C tends to dominate the thinking of even postmodern educators. Too much is lost when new ideas are translated into those commonsense formulas, and so the practitioner-oriented literature keeps getting farther out of contact with what is going on in the underlying behavioral sciences.

A more immediate reason for learning to think in connectionist terms is in order to make progress in dealing with what is now vaguely referred to as tacit knowledge. Business people have suddenly become intensely aware of it and it is much discussed in the knowledge management literature (e.g., Nonaka & Takeushi, 1995; Stewart, 1996). It is knowledge that isn't written down anywhere. It tends to vanish when people leave or when working groups dissolve. It isn't private facts. It is the skills and what I will discuss in Chapter 5 as 'implicit understanding' that make people and organizations smart. For folk theory of mind it will remain forever mysterious, because it cannot be pinned down to objects in the mind. And so it also tends to be ignored in education. I have mentioned number sense as an example, but there is more: historical sense, geographical sense, moral sense, tact, taste and appreciation, background knowledge (what you need, for instance, in order to understand a Jane Austen novel in the context of the time and place and social world of its characters and of Austen herself as a part of that world), and acquired instincts —"a nose for news," the ability to recognize promising leads, problems, or ideas in a domain (Bereiter & Scardamalia, 1993, Ch. 5). Educators readily acknowledge these as things that really count, that may be of more enduring value than any particular skills or items of knowledge, but they are addressed only through hand-waving and wishful thinking.

Educators and knowledge managers alike need a way of thinking that brings these kinds of knowledge into the natural order. Folk theory and the psychologies grounded in it are out of their depth here. Recognizing that an alternative theory is only starting to take shape, however, what should those of us concerned with education and other knowledge arts do in the meantime? The following are suggestions:

- *Minimize cognitive speculation*. We need mentalistic description but we should not go farther with it than necessary. In Chapter 4 I will try to show how far we can get in dealing with problems of understanding without having to speculate on what is in the mind of the understander.
- *Recognize that talk about mental content and mental operations may be only a manner of speaking, useful for some purposes and not others.* It may turn out to be more

than a manner of speaking, but it is never going to turn out to be less. And as far as I can see, there is no practical loss and considerable practical gain from treating mental content and mental operations as useful fictions.

- *Instead of hypothesizing what does or does not come naturally, look at the evidence on what does or does not profit from instruction and on kinds of attainments that are particularly easy or difficult.* We know that many students who master whole-number arithmetic get lost in ratios and proportions. Maybe this has something to do with how our brains are built, and it would be interesting to know; but in the meantime it stands as a brute fact that responsible mathematics educators must contend with. We also know that reading comes fairly easily to about 85 percent of children, regardless of how they are taught, and that the other 15 percent have difficulties, sometimes very serious ones. None of the argument about whether reading is a natural outgrowth of innate language competence or a kind of unnatural act alters the proportions one bit. The problem is exacerbated, however, by ideologically committed reading specialists, some of whom ignore the 85 percent and some of whom ignore the 15 percent.
- *Don't worry about distinguishing knowledge from other kinds of competence. If it is mental and useful, call it knowledge.* Jane has studied weather maps, the location of the jet stream, and has compared the predictions from several weather stations, from which she concludes that it is going to be sunny tomorrow. Joe believes it is going to rain because it always rains on their picnic. Lana has a good feeling about tomorrow and is sure it will be a nice day. Hugh feels in his bones that rainy weather is coming. Do some of these have knowledge and others not? Does the answer depend on who turns out to be right? Or would you say that no one *knows* what tomorrow's weather will be? For practical purposes, we are better off avoiding such questions than trying to answer them. As I will try to show in Chapter 5, there is value in distinguishing among the different *kinds* of knowledge represented in these examples, but in practical terms they all represent knowledge that is relevant to action. Lana's feelings may not be good for

predicting the weather but in other situations they may be a more reliable guide than Jane's bookish analyses and Joe's experience-based cautiousness. They are all kinds of knowledge worth cultivating, worth learning how to use.

These suggestions add up to a restrained kind of mentalism. We should not, like behaviorists, avoid mentalistic descriptions. We cannot function in the knowledge arts without talking about what people think, feel, know, and understand. Indeed, we want to widen not narrow the range of what counts as part of human knowing. But in order to do this sensibly, we need to avoid premature commitments as to the nature of what goes on between the ears. That does not mean being theoretically unmindful. It means trying to adopt a course that will not prove gravely wrong however the theoretical cognitive issues are eventually worked out.

Conclusion

I am always dubious about educational discussions having to do with the brain. Too often they consist of excited reports of advances in neuroscience, accompanied by hints that any educator who fails to keep up with them will soon be out to pasture—but with no clear indication of why educators in general should be more than casually interested. Is this chapter further ado about matters that can safely be put on the 'read later' pile? Of course I think the answer is no, but the reasons I can advance at this point cannot be very compelling. I must trust that conviction will grow as ideas are developed further and in more educationally targeted ways in later chapters.

No law requires that a theory of mind be biologically plausible. The endurance of the existing theory of mind is proof of that. To recapitulate an earlier argument, we need a mentalistic level of description, a way of talking about behavior in terms of intentions, beliefs, and the like. But we could employ mentalistic description without a concept of mind. That would seem to be what people do when they use mentalistic terms to explain the behavior of a computer or a fish. The concept of mind enters when we try to give more complete and integrated explanations of behavior. But mind could still be a hypothetical construct with no necessary correspondence to biological reality. It could be like the *market* , as referred to by financial reporters: "The market responded quickly to news of the assassination," "The market took a wait-and-see attitude toward the offering," etc. Although 'the market' is endowed with humanlike qualities, it has no necessary correspondence to anything in the material world. It is a way of expressing general trends and tendencies. Perhaps that is all 'the mind' is required to be.

I have introduced two reasons, however, why a theory of mind for education needs to be more realistic. One is in order to deal with the vast areas of learning that do not lend themselves to formulation as mental content but that can better be understood as mental adaptations to things in the world. The other is in order to deal with what the newborn’s brain brings into the world with it: human nature, instinct, innate knowledge and dispositions—however we may choose to formulate it. I will not be pursuing this second reason further in this book. I expect that the rediscovery of human nature is going to have much significance for education, but I am in no position to argue that case at present or to contribute to its furtherance. But I will be pursuing the first reason throughout the remaining chapters. I think I can show that a biologically more realistic theory of mind is better able to handle most of the major goals of education, including especially teaching for understanding.

Perhaps the main reason why we need a more realistic theory of mind for the knowledge age, however, is in order to get the mind disentangled from knowledge. Obviously mind and knowledge are related. But if, as some economists say, the main wealth-generating activity of the future is going to be knowledge production, then it seems that two things are required: (1) to conceive of knowledge as something other than stuff inside individual people's minds and (2) to understand the role of individual minds in societal knowledge production. Folk theory is not up to either of these requirements. A connectionist view of mind makes the first conceivable and provides a least a starting point for the second. In the next chapters I will try to move beyond this starting point.

---

[1] The usual disclaimers apply, the likening of the brain to a connectionist network running on a computer being only a crude analogy (Smolensky, 1988). The point I am trying to make can survive the crudity, I believe, provided the reader is willing to put up with a style of exposition that treats brains as if they really are connectionist networks.

[2] I feel that I am belaboring the obvious here. The reason for doing so is that I have encountered several sophisticated colleagues who miss the point, being stuck on insisting that connectionist networks contain rules, just different kinds of rules, from more traditional artificial intelligence programs.

[3] Since performing this little experiment, I have acquired a new version of the word processor with, evidently, an improved spelling checker. It suggested “psychology” in all cases, except for “cychology,” which it still suggested was “cytology.” Further experimentation might lead to plausible guesses about the content of the new or revised rules that accounted for the improvement.

[4] The number of connections this would mean with a dictionary of, say, 30,000 words suggests why this approach to spell checking has not caught on yet.

[5] That is the number that Hanna, Hanna, Hodges, and Rudorf (1966) arrived at in an early attempt to work out rules for English spellings. However, even with that many rules a computer simulation was still able to spell correctly only about half of the 17,000 most common words (Simon & Simon, 1973).

[6] Hebb and Thompson (1954) argued thus, offering evidence that the smarter the species, the more emotionally unstable it is, the apparent pinnacle having been reached by the chimpanzees. Anyone who has worked around these genetic neighbors of ours knows that their portrayal in movies as fun-loving free spirits is achieved only by editing out all their insane rages and panics. Hebb and Thompson went on to suggest that the trend toward emotional instability may not have reversed when it got to homo sapiens, that we may be the most unstable species of all, except for the fact that we are able to arrange our physical and social environment in ways that keep us sane most of the time.

[7] Evidence of these innate dispositions and capacities is gathered in Hirschfield and Gelman (1994).

[8] The older connectionism did not entail rejecting the mind-as-container metaphor. Its distinctiveness lay in proposing that the items in the mental container were little bits of ideas, which acquired structure and force through connections to one another that were formed through experience. Thus the older connectionism is really more similar to some symbolic artificial intelligence models (e.g., Anderson, 1983) than it is to the new connectionism.

[9] This may prove to be a tactical error. Tribalism and its accompaniment, tribal warfare, are still to be found in the social sciences and are especially common in education. So, to the extent that using the term ‘connectionism’ identifies me with a tribe, I can expect hostility from that tribe’s enemies.